



JRC SCIENTIFIC INFORMATION SYSTEMS AND DATABASES

AAXY User Guide

Documentation
Version 1.0beta

Martino Pesaresi

Luca Maffenini

2016

This publication is a Scientific Information Systems and Databases report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Technical Support

For technical support send an email to ghsl-tools@jrc.ec.europa.eu

JRC Science Hub

<https://ec.europa.eu/jrc>

JRC102959

PDF ISBN 978-92-79-61818-5 doi: 10.2788/782191

Luxembourg: Publications Office of the European Union, 2016

© European Union, 2016

Reproduction is authorised provided the source is acknowledged.

How to cite: Martino Pesaresi, Luca Maffenini; AAXY User Guide; doi: 10.2788/782191

All images © European Union 2016

Table of contents

Abstract	1
1 Introduction	2
2 Installing AAXY	3
2.1 System Requirements	3
2.2 Installing Matlab Runtime	3
2.3 Installing AAXY	3
2.4 Uninstalling AAXY	3
3 Using the AAXY command line interface	4
3.1 Supported Image files	4
3.2 Start the software and check the help	4
3.3 Basic ENDI extraction	5
3.4 Advanced options	6
3.4.1 Change output ENDI filename	6
3.4.2 Select input file descriptors	6
3.4.3 Select positive reference values	7
3.4.4 Select negative reference values	7
3.4.5 Choose type of ENDI measure	8
4 Troubleshooting	9
4.1 AAXY error messages	9
4.2 Matlab and GDAL errors	12
References	13
List of abbreviations and definitions	14

Abstract

AAXY stands for Associative Analysis between X and Y, it's a special analytic method developed in the frame of the Global Human Settlement Layer (GHSL) project and it's being used for extraction of information from satellite images.

This software allows user to extract information (Evidence-based Normalized Differential Index or ENDI measure) from satellite data and it's been used for the production of the GHSL products such as GHSL BUILT and GHSL POP.

More details about the measure, the method and its actual applications can be found in the papers in the Reference section.

This guide provides instructions about installing and using the AAXY command line tool on a Windows computer.

1 Introduction

AAXY is a command line tool that computes ENDI (Evidence-based Normalized Differential Index) between an input and a reference geodata using Associative Analysis. For more details about ENDI measure and Associative Analysis refer to the papers in the Reference section.

Given a raster geodata in INPUT (x) and a binary REFERENCE raster (y) it produces a new raster with ENDI values (range -1:1) and same geographic extension.

The tool is developed using Matlab and it's now available as Windows executable.

2 Installing AAXY

2.1 System Requirements

To use AAXY your system must meet the following minimum requirements:

Operating system	Processor	Disk space	RAM
Windows 7 or Windows 10	Any Intel or AMD x86-64 processor	25 Mb (+ 700 Mb for Matlab Runtime)	2 GB

To have AAXY running on your computer you need to:

1. Install Matlab Runtime
2. Install AAXY

2.2 Installing Matlab Runtime

Verify the MATLAB Runtime is installed and ensure you have installed version **9.0 (R2015b)**.

If not already installed, download the Windows 64-bit version of the MATLAB Runtime for R2015b from the MathWorks Web site:

<http://www.mathworks.com/products/compiler/mcr/index.html>

For more information about the MATLAB Runtime and the MATLAB Runtime installer, see Package and Distribute in the MATLAB Compiler documentation in the MathWorks Documentation Center.

2.3 Installing AAXY

AAXY is a portable command line application. There is no Windows installer, you just need to download the executable from the European Commission website and save it on your hard drive.

The executable can be downloaded from the GHSL project webpage:

<http://ghsl.jrc.ec.europa.eu>

2.4 Uninstalling AAXY

To remove AAXY just delete the executable file itself and uninstall the Matlab Runtime.

3 Using the AAXY command line interface

3.1 Supported Image files

AAXY uses GDAL to read and write georeferenced raster images and it supports the following raster formats:

Format	Official Website
GeoTIFF	https://trac.osgeo.org/geotiff
JPEG2000	https://jpeg.org/jpeg2000

Even if GDAL supports more raster formats, only these have been actually tested. Using other GDAL supported formats is possible but not recommended.

For the complete list of such formats check the official GDAL documentation:

http://www.gdal.org/formats_list.html

3.2 Start the software and check the help

1. Open the Command Prompt or the PowerShell

Open the Command Prompt window by clicking the **Start** button. Picture of the Start button, clicking **All Programs**, clicking **Accessories**, and then clicking **Command Prompt**.

```
Microsoft Windows [Version 10.0.10586]
(c) 2015 Microsoft Corporation. All rights reserved.

C:\Users\maffelu.ISIS>
```

2. Navigate to the directory containing the aaxy.exe file (e.g. Downloads)

```
C:\Users\maffelu.ISIS>cd Downloads
```

3. Type `aaxy.exe` and press enter

```
C:\Users\maffelu.ISIS\Downloads> aaxy.exe
```

You should see the help text, which confirms the software works properly.

Otherwise you should receive an error. In this case double check that:

- Matlab Runtime is correctly installed
- your computer meets the minimum system requirements

3.3 Basic ENDI extraction

AAXY allows you to extract ENDI information from geographic raster data.

Given:

- a georeferenced raster **input.tif** (containing the data to process)
- a georeferenced raster **reference.tif** (containing the reference information)
- a quantization parameter value

it will produce a georeferenced raster **input_endi.tif** which contains the ENDI information.

Open the Command Prompt and type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10
```

The two rasters input.tif and reference.tif have to overlap, both in geographic projection and extension as well as in pixel resolution.

Raster input.tif can be a single or multi descriptor (a.k.a band) file.

Raster reference.tif has to be a single descriptor file, with binary or multi values content.

In case of binary reference data:

- value 1 is for pixels which represent the positive examples of the target
- value 0 is for pixels which represent the negative examples of the target

In case of multi value reference data:

- all positive values represent the positive examples of the target
- value 0 represents the negative examples of the target

This default behaviour can be changed using the advanced options.

The quantization parameter is an integer value that represents the uniform quantization factor applied to the input.tif data.

3.4 Advanced options

3.4.1 Change output ENDI filename

You can change the default output ENDI filename and path.

To save the output in the same folder with a different name type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -o my_endi_custom.tif
```

To save it in another directory specify the new filename with the complete path:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -o D:\new_path\my_endi_custom.tif
```

3.4.2 Select input file descriptors

When you have a multi descriptor input raster, all descriptors are used by default, but you can choose to use just one or few of them.

Let's say that input.tif has 4 descriptors, to use only the first one type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -dx 1
```

To use descriptors 3 and 4 type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -dx 3,4
```

To specify a different quantization for each descriptors type:

```
> aaxy.exe -x input.tif -y reference.tif -q 8,16 -dx 3,4
```

3.4.3 Select positive reference values

When you have a multivalue reference raster (a.k.a a raster with values from 0 to N number of classes), you can specify to use a single class or multiple class as positive reference.

Let's say that reference.tif has 7 classes (values go from 0 to 6), to use only class 2 as positive example type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -py 2
```

To use values 1, 3 and 4 type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -py 1,3,4
```

If this option is not specified, all positive classes are used as positive examples.

3.4.4 Select negative reference values

When you have a multivalue reference raster (a.k.a a raster with values from 0 to N number of classes), you can specify to use a single class or multiple class as positive reference.

Let's say that reference.tif has 7 classes (values go from 0 to 6), to use only class 3 as negative example type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -ny 3
```

To use values 3, 5 and 6 type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -py 3,5,6
```

If this option is not specified, class with value 0 is used as negative example.

3.4.5 Choose type of ENDI measure

You can choose to compute different types of ENDI measure:

1. **ENDI absolute:** $\Phi_E^a = \frac{f_{pos} - f_{neg}}{f_{pos} + f_{neg}}$

where f_{pos} and f_{neg} are the frequencies of the joint occurrences among X (antecedent) data instances and the positive and negative Y (consequent) reference instances, respectively.

2. **ENDI relative:** $\Phi_E^b = \frac{p_{pos} - p_{neg}}{p_{pos} + p_{neg}}$

with the empirical probabilities p_{pos} and p_{neg} calculated as $p_{pos} = \frac{f_{pos}}{N_{pos}}$ and $p_{neg} = \frac{f_{neg}}{N_{neg}}$,

with N_{pos} , N_{neg} denoting, respectively, the numerosity of the positive and negative training samples.

3. **ENDI average:** $\Phi_E^{ab} = \frac{\Phi_E^a + \Phi_E^b}{2}$

is the mean value between the other 2 ENDI measures.

To compute ENDI absolute type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -e abs
```

To compute ENDI relative type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -e rel
```

To compute ENDI average type:

```
> aaxy.exe -x input.tif -y reference.tif -q 10 -e avg
```

Default value is to compute ENDI absolute Φ_E^a .

4 Troubleshooting

4.1 AAXY error messages

Missing mandatory x or y input! Type "aaxy.exe --help" for support

Missing mandatory parameter '-y'. Type "aaxy.exe --help" for support

Missing mandatory parameter '-x'. Type "aaxy.exe --help" for support

Missing mandatory parameter '-q'. Type "aaxy.exe --help" for support

X or Y or Q inputs are missing from the command line. You need to specify at least -x, -y and -q options.

The value of '-x' is invalid. Input file "landsat.tif" does not exist!

The file *landsat.tif* can't be found in the current folder. Check the file exists and specify the complete path in case it's in another folder.

The value of '-x' is invalid. File "unsupported.avi" not recognized as supported file format

The file *unsupported.avi* is not supported as input file. Convert it to a supported format like GeoTIFF.

The value of '-y' is invalid. Input file "reference.tif" does not exist!

The file *reference.tif* can't be found in the current folder. Check the file exists and specify the complete path in case it's in another folder.

The value of '-x' or '-y' is invalid. File "document.pdf" not recognized as supported file format

The file *document.pdf* is not supported raster file. For a list of supported file type check the *Supported Image files* section in this manual.

Missing or empty value for some parameter

You specified a parameter (like -x or -q) with an empty value. Remove the parameter or add a value.

The value of '-q' is invalid. Expected input to be a vector.

You specified a not vector nor scalar numeric value.

Found 2 quantization values for 3 descriptors. Choose 1 global scalar value or 1 value per descriptor

The number of values for -q and -dx parameters doesn't match.

The value of '-o' is invalid. Directory "X" does not exist!

The specified directory "X" doesn't exist, double check the path is correct.

The value of '-dx' is invalid. Expected input to be a vector.

Double check that the number specified is a positive integer or a vector in the form 1,2,3.

The value of '-dx' is invalid. Expected input to be an array with all of the values > 0.

The descriptors can only be integer and positive. Allowed values go from 1 to the max number of descriptor in the input image.

The value of '-dx' is invalid. Expected input to be an array with all of the values <= 3.

The descriptors contain a value that's too high.

The value of "-dx" is invalid, max number of elements is 3

You specified too many descriptors values.

X and Y have not the same resolution. X = 102x120, Y = 868x947

X and Y inputs must have the same pixel resolution. Cut or resize the images.

X and Y have not same projection or extension

X and Y contains geographic information that doesn't match. They should represent the same geographic location, have the same geographic extension and be represented in the same projection.

The value of '-py' is invalid. Expected input to be a vector.

Double check that the number specified is a positive integer or a vector in the form 1,2,3.

The value of '-py' is invalid. Expected input to be an array with all of the values > 0.

Double check all values are positive.

The value of '-py' is invalid. Expected input to be an array with all of the values <= 6.

Double check all values are less then the maximum for the specific reference input.

The value of '-ny' is invalid. Expected input to be a vector.

Double check that the number specified is an integer greater or equal to 0 or a vector in the form 0,1,2.

The value of '-ny' is invalid. Expected input to be an array with all of the values ≥ 0 .

Double check all values are greater or equal to 0.

The value of '-ny' is invalid. Expected input to be an array with all of the values ≤ 6 .

Double check all values are less than the maximum for the specific reference input.

The value of '-py' or '-ny' are invalid. Value 2 can not be used in both positive and negative examples

The specified number 2 has been set in both positive and negative examples (e.g. -py is 1,2,3 and -ny is 0,2; 2 is in both and it can not be).

The value of '-e' is invalid. Expected input to match one of these strings:

'abs', 'rel', 'avg'

The input, X, did not match any of the valid strings.

The choice on ENDI computation is limited to three options "abs", "rel", "avg".

Replace "X" with one of these three values.

4.2 Matlab and GDAL errors

There could be other low lever errors due to Matlab or GDAL, like issuer dealing with reading or writing unsupported raster files.

In such case please refer to the documentation available at the official websites:

Matlab	https://uk.mathworks.com/help/matlab/index.html
GDAL	https://trac.osgeo.org/gdal/wiki/FAQ

References

AAXY is based on the studies performed at the Joint Research Center (JRC) in Ispra, Italy.

For more details about ENDI and the Associative Analysis techniques used check:

A New Method for Earth Observation Data Analytics Based on Symbolic Machine Learning

Pesaresi, M.; Syrris, V. & Julea, A.

Remote Sensing, 2016 Remote Sens. 2016, 8(5), 399;

doi:10.3390/rs8050399

<http://www.mdpi.com/2072-4292/8/5/399>

Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014

Pesaresi, M.; Ehrlich, D.; Ferri, S.; Florczyk, A.; Freire, S.; Halkia, S.; Julea, A.; Kemper, T.; Soille, P. & Syrris, V.

JRC Technical Report EUR 27741 EN;

doi:10.2788/253582

<http://publications.jrc.ec.europa.eu/repository/handle/JRC97705>

Benchmarking of the Symbolic Machine Learning classifier with state of the art image classification methods

Pesaresi, M.; Syrris, V. & Julea, A.

JRC Technical Report EUR 27518 EN;

doi:10.2788/638672

<http://publications.jrc.ec.europa.eu/repository/handle/JRC97964>

GHSL Project

<http://ghsl.jrc.ec.europa.eu>

External components and libraries:

MATLAB Runtime

<http://it.mathworks.com/products/compiler/mcr>

GDAL

<http://www.gdal.org>

List of abbreviations and definitions

AAXY

Name of the software described in this user guide. It stands for Associative Analysis between X and Y.

ENDI

Evidence-based Normalized Differential Index, a kind of interestingness metric measure.

raster

An image composed of a complete grid of pixels.

geodata

An image that has geographic information embedded in the file, like GeoTIFF.

GDAL

A computer software library for reading and writing raster and vector geospatial data formats.

GeoTIFF

A public domain metadata standard which allows georeferencing information to be embedded within a TIFF file.

JPEG2000

An image compression standard and coding system.

descriptor

A layer of the input raster data. In a multiband image each band is a descriptor.

Europe Direct is a service to help you find answers to your questions about the European Union
Free phone number (*): 00 800 6 7 8 9 10 11
(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server <http://europa.eu>

How to obtain EU publications

Our publications are available from EU Bookshop (<http://bookshop.europa.eu>),
where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.
You can obtain their contact details by sending a fax to (352) 29 29-42758.

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



EU Science Hub

ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub

